Statistical Modeling of Shallow Velocity Heterogeneities with Validation against Strong Ground Motion

1/1/2012 – 12/31/2013

Principal Investigator*

Kim B. Olsen,
Dept of Geological Sciences, San Diego State University
5500 Campanile Dr., San Diego, CA 92182-2010
ph 619 594 2649
email: kbolsen@mail.sdsu.edu

* W. Savran, B. H. Jacobsen, and K. Withers contributed significantly to the work reported here.

ABSTRACT

State-of-the-art velocity models poorly resolve the meter-scale structure of the earth's crust, which play an important role in ground motion amplification as frequencies increase. Here, we analyze borehole velocity logs and shallow shear-wave velocity (V_s30) data in the Los Angeles basin to construct a statistical model of the small-scale heterogeneities for ground motion prediction purposes. Our analysis suggests that a von Karman distribution with Hurst numbers (ν) of 0.0 - 0.1, correlation lengths (a) of 50-150 m, and standard deviations from the mean fluctuations around 5% characterize the data best. To quantify and validate the effects of the statistical distribution on ground motions we use finite-difference solutions to the 3D wave equation including frequency-dependent attenuation. We simulate 0 - 2.5 Hz linear viscoelastic waves through 3D velocities structures derived from the Southern California Earthquake Center (SCEC) Community Velocity Model (CVM) SI-4.26 with a minimum S-wave velocity (V_s) of 200 m/s. Our analysis shows that statistical distributions of small-scale heterogeneities with parameters constrained from data can amplify or de-amplify ground motions by up to a factor of two. However, small-scale scattering included in the wave propagation for the 2008 $M_w5.4$ Chino Hills, CA, earthquake, by superimposing the statistical models on the velocity model, improves the goodness-of-fit (GOF) between data and synthetics by only ~5%. A relation between the quality factor (Q_s) and V_s of Q_s/V_s~100-150 (V_s in km/s) provides the largest GOF values and best-fitting peak motion attenuation with distance when accounting for a scattering medium. Furthermore, we show that the majority of the scattering recorded in ground motions originates as path effects as waves propagate through the basins, as compared to local site-specific scattering. We also find lower-velocity sediments and the deep crust contribute approximately equally to the strength of the scattering recorded in ground motion records.

The effects of the statistical distributions on the synthetic ground motions for the deep (~14 km) Chino Hills earthquake consist of highly incoherent and scattered amplifications and de-amplifications. On the other hand, we find that shallow sources located on the boundary to a sedimentary basin can generate bands of strong amplification aligned in the direction of the ray paths. The nature of these bands depends strongly on the incidence angle of the waves into the sediments. Moreover, this banded amplification pattern is absent for sources deeper than 1-2 km. These results imply that surface rupture on a range-bounding fault may generate different patterns of ground motion shaking along lines parallel to the fault as compared to profiles perpendicular to the fault.

Our analysis shows that the statistical distributions of small-scale heterogeneities have only secondary effects on the misfit between data and synthetics for the Chino Hills earthquake. However, incorporating a V_s30-based geotechnical layer brings the ground motion metrics from the synthetics averaged with distance from the source closer to those from data by up to 20-50%. The largest overall improvement in GOF is obtained by increasing the moment for the Chino Hills event from $M_w5.4$ to 5.5, an adjustment within the uncertainty of the magnitude of the earthquake.

Bibliography of Publications Resulting from the Work Performed Under the Award

Savran, W.H. and K.B. Olsen (2014). Deterministic simulation of the Mw5.4 Chino Hills event with frequency-dependent attenuation, heterogeneous velocity structure and realistic source model, *Seism. Res. Lett*. **85:2**, 498.

Savran, W.H., K. B. Olsen, and B. H. Jacobsen (2013). Statistics of Velocity Structure in the Los Angeles Area, Procs of the American Geophysical Union Fall Mtg, San Francisco, December 2013 (S23A-2481).

Withers, K., K.B. Olsen, and S.M. Day (2013). Deterministic high-frequency ground motion using dynamic rupture along rough faults, small-scale heterogeneities, and frequency-dependent attenuation, Procs of the American Geophysical Union Fall Mtg, San Francisco, December 2013 (S52C-08).

Final Technical Report

INTRODUCTION

Gaining a solid understanding of the expected range of ground motions from future large earthquakes is imperative so structures can be properly engineered to withstand this shaking. Local site effects can drastically amplify the ground motion from an earthquake, with prominent examples from the 1989 Loma Prieta earthquake in the Marina District of San Francisco and the 1985 Michoacan earthquake in Mexico City. As the capabilities of deterministic seismic wave propagation advance, characterizing the small-scale variation (from tens to hundreds of meters) of the near surface soil deposits becomes necessary in modeling site-specific amplifications. The variability in the site amplification is important for the design of lifelines such as bridges and pipelines, as these structures extend over considerable length parallel to the ground. Current velocity models (e.g. the Southern California Earthquake Center (SCEC) Community Velocity Model (CVM)-S 4.0 (Magistrale et al., 2000; Kohler et al., 2003), CVM-H 11.9 (Süss and Shaw, 2003), CVM-SI 4.26 (Lee et al., 2014) insufficiently resolve the small-scale variability known to exist in the earth's crust.

Due to data acquisition costs, it is not feasible to resolve the small-scale structure using direct measurements. Instead, the small-scale heterogeneities may be described by statistical distributions, such as von Karman auto covariance functions (1.1) with Fourier transform (1.2)

$$\Phi_{v,L}(r) = \sigma^2 \frac{2^{1-v}}{\Gamma(v)} \left(\frac{r}{a}\right)^v K_v\left(\frac{r}{a}\right) \tag{1.1}$$

$$P(k) = \frac{\sigma^2 (2\sqrt{\pi}a)^E \Gamma(v + E/2)}{\Gamma(v)(1 + k^2 a^2)^{v+E/2}}, \tag{1.2}$$

(Tatarski, 1961), as well as Gaussian auto correlation functions. Constraints on the parameters of the statistical parameters - the correlation length, *a,* Hurst exponent, ν, and standard deviation, σ - were obtained from sonic logs and digitized geological maps (e.g., Wu et al., 1994; Levander et al., 1994; De et al., 1994; Holliger and Levander, 1992; Holliger, 1996; Holliger, 1997; Dolan et al., 1998). These studies report Hurst exponents of 0.0-0.3 and correlation lengths between 60-160 m in the vertical direction (Holliger 1996), and horizontal to vertical anisotropy (H/V) between ~2-5 (Holliger and Levander, 1992; Wu et al., 1994). We note that 'fractal dimension' is sometimes used instead of Hurst exponent. Higher fractal dimensions represent more complexity in the detail of a fractal with a change in scale. It can be shown that D = E+1-ν, where E is the Euclidian dimension, i.e., a high fractal dimension corresponds to a low Hurst exponent.

Other studies have assessed the statistical parameters based on wave propagation simulations. Frankel and Clayton (1986) and Hartzell et al. (2010) used ν=0.0 with a=5-10 km and σ=5-10%. Olsen and Jacobsen included statistical distributions with ν=[0.5-0.5], a=250 m and σ=5-10%. Imperatori and Mai (2013) used a Hurst exponent of 0.3 with correlation lengths on the order of hundreds of meters to kilometers. Frankel and Clayton (1986) generated small-scale heterogeneities with three distinct autocorrelation functions, Gaussian, exponential (corresponding to ν = 0.5), and self-similar (corresponding to a ν = 0.0). By testing end-member cases of the von Karman autocorrelation function Frankel and Clayton (1986) provided constraints on ν estimates for the small-scale heterogeneities that fall in the range [0.0, 0.5] with a preferred value of 0.0. There is considerable variation in the autocorrelations used among the studies, reflecting the variable constraints from the underlying data. For example, the correlation lengths suggested from Frankel and Clayton (>10 km) were obtained using teleseismic sources, which provide limited constraints at near-field stations. Here, we validate the resulting statistical distributions of small-scale heterogeneities against strong motion data using 3D finite difference simulation of wave propagation.

The simulation studies mentioned above focused on crustal-scale models, and in part due to computational limitations, did not include near-surface low-velocity material (LVM). Ignoring the near-surface LVM in these simulations was justified in part by the omission of anelastic attenuation. However, the presence of the LVM, the associated scattering, and attenuation parameters are critical for ground motion estimation. Previous studies have estimated intrinsic attenuation for S-waves (Q_s) and P-waves (Q_p) by matching simulations to strong motion data (e.g., Olsen et al., 2003; Graves and Pitarka, 2010; Taborda and Bielak, 2013), where Q_s is taken as a function of the local V_s, and Q_p is proportional to Q_s. However, these $Q_s -V_s$ relations may be biased by different (artificially imposed) minimum velocities for the sediments (often dictated by computational limitations). Moreover, by including small-scale velocity heterogeneities in the LVM, the associated scattering attenuation may redefine the parameters for intrinsic attenuation, which are usually estimated using smooth velocity models.

In this study, we use a data set of shear-wave velocities averaged over the top 30 meters (V_s30) and 38 P-wave (V_p) sonic logs to determine the parameters characterizing the small-scale velocity fluctuations in the Los Angeles Basin. We validate the results

against strong-motion data from the 2008 M_w5.4 Chino Hills earthquake using 3D wave propagation simulation and engineering relevant quantitative goodness-of-fit (GOF) metrics based on Olsen and Mayhew (2010). This validation includes determining the values of Q_s and Q_p in the presence of scattering attenuation by small-scale heterogeneities. We then investigate the relative importance of seismic scattering from localized site effects within the sedimentary basin and along the entire path from source to surface record. Then, we examine the depth dependency on ground shaking caused by seismic scattering, as well as the nature of scattering for shallow versus deep seismic sources and their proximity relative to sedimentary basins. Finally, we analyze the improvements in GOF values from of adding a shallow geotechnical layer (GTL) to the velocity model as well as adjusting the moment of the event with a reasonable amount.

SPATIAL STATISTICS OF LOS ANGELES BASIN

A primary goal of this study is to understand the spatial statistics of the variation in the seismic velocity structure in order to accurately model strong ground motion from earthquakes. Toward this goal we use variogram analysis on three separate data sets consisting of V_s30 measurements and borehole sonic logs, to estimate the parameters to model the small-scale variability for von Karman autocorrelation functions (Hurst exponent, standard deviations, and correlation length).

V_s30 Measurements

First, we analyze the horizontal variation of near-surface small-scale heterogeneities in greater Los Angeles. We compiled 639 V_s30 measurements from several sources shown in Fig. 1. We used 350 measurements incorporated in the SCEC CVM-S 4.0 (Magistrale et al., 2000), 102 measurements from (Louie, 2005; 2007), and 187 measurements from Yong et al. (2013).

V_s30 is defined as

$$Vs_{30} = \frac{30}{\sum \frac{d}{Vs}},$$

(1.3)

where d is the layer thickness from the 1D profile below the site of interest, and V_s is the depth-dependent S-wave velocity. We analyze the V_s30 data in the time-domain using the classical estimate of semi-variance given by Matheron (1963) in Eq. 1.4.

$$\tilde{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]$$

(1.4)

This variogram estimator has a vetted history in geostatistics since its derivation and is frequently used in kriging algorithms where developing a model of spatial correlation is paramount. The factor of 2 appears in the denominator accounting for the double

representation of field values (x_1, x_2) and (x_2, x_1) in the summation. The estimate requires binning of the data to specify lag distances, and this estimate performs the best when there are near-equal representations of the data in each bin ($N(h)$).

We can directly interpret the correlation length of the velocity structure given an estimate of the variogram, while we fit a model to the variogram to determine the Hurst exponent. Here, we model the variogram using a power-law or fractional Brownian motion (fBm) form (Li and Lake, 2010; Mela and Louie, 2001), based on the results from previous studies (e.g. Wu et al., 1994; Holliger, 1996; Frankel and Clayton, 1986), see Eq. 1.5. The factor of 2 appears in the exponent because the autocorrelation scales as the square of the amplitude. Following Eq. 1.5, the Hurst exponent can be estimated as the half-slope of the linear regression in logarithmic space.

$$\tilde{\gamma}(h) = ch^{2\nu}$$

(1.5)

For the V_s30 measurements, the data are sparse and irregularly located. To account for this we assign a relatively large bin spacing of the variogram of 500 m. We restrict the V_s30 measurements to a 40 km x 40 km region located inside Los Angeles basin (resulting in 292 values used in the analysis) to maintain consistency with the sonic log analysis (see the following section). Also, to obtain a similar representation at all lag distances, we restrict the maximum lag to 30 km. Fig. 2 shows the horizontal variogram calculated from all of the analyzed V_s30 data. We interpret the correlation length of the heterogeneity in the V_s30 values to be approximately 5-10 km (where the variogram levels out), and the Hurst exponent to be $\nu = 0.33$. This analysis provides an understanding of the correlation length of larger-scale basin features, expected to be captured by and included in the SCEC CVMs.

While many of the V_s30 values are measured directly in boreholes, a significant number are obtained using inversion of surface wave data. The latter estimate tends to smooth the actual velocities, potentially resulting in biased (high) Hurst exponents, and (low) fractal dimensions. We therefore consider the Hurst exponent and correlation length estimated from the V_s30 values as an upper bound on the estimates of the true ν. While these V_s30 values do not provide enough small-scale resolution to accurately model the fine-scale spatial statistics present in Los Angeles basin they are consistent with Hurst exponents estimated from Levander et al. (1994) and the correlation lengths suggested by Frankel and Clayton (1986).

Analysis of Borehole Logs

In order to constrain the small-scale structure in Los Angeles basin, we require much more densely sampled data than that provided by V_s30 measurements, as discussed above. Velocity logs provide regularly and densely sampled direct measurements of in situ seismic velocities. The following section contains our analysis of velocity logs from the Los Angeles basin.

We introduce an alternative variogram estimate (moving window formulation) derived by Li and Lake (2010) to analyze the sonic log measurements. The moving-window formulation (Eq. 1.6) provides an estimate utilizing all available data pairs at all lag distances, whereas other variogram estimates (such as the Matheron, 1963, method discussed above) provide less data pairs as observation distance increases. We use the moving window method for the borehole logs as it reduces the variance at large lag distances, but does not necessarily provide a more accurate estimate as compared to the Matheron (1963) estimate.

$$\hat{\gamma}_{N1}(h) = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{1}{2m}\sum_{D\in i,h}[Z(x_i) - Z(x_j)]^2\right\} . \qquad (1.6)$$

The moving-window variogram method is incorporated into an inversion routine to perform a search of the parameter space minimizing the chi-squared test for GOF between the model and the data in the deep borehole logs. Here, we assume that each borehole log is a realization of a single random field of velocity fluctuations representing the entire Los Angeles basin, and the statistics are stationary. These assumptions permit us to average the variograms together to reduce the variance, creating an expected variogram estimate for all of the deep borehole logs. This expected variogram represents the observations used in our chi-square test.

We pre-process each borehole log individually. First, we remove the depth-dependent mean using a long-period median filter, which leaves us with the fluctuations around the mean. Then we use median filters with different window sizes to remove long-period trends (window length 100 m) as well as artifacts resulting from the down-hole logging procedure (window length 5 m). We assume that the large spikes seen in the measurements are due to measurement errors resulting from either decoupling with the logging tool and the host rock or cycle skips, a phenomenon resulting in anomalously high transit times or low velocities. Fig. 3 shows the preprocessing method performed on one borehole log, namely apal1 located approximately 5 km north of Long Beach, CA, along with the moving-window variogram (Fig. 3c) estimate for this single borehole.

Deep Borehole Measurements

We acquired 38 borehole V_p sonic logs (Personal Communication, Shaw and Plesch, 2012) located throughout the Los Angeles basin, extending to more than 3 km depth and containing more than 300,000 measurements. Fig. 4 shows the locations of the boreholes. Some logs are missing sections of data, and these values are removed from the resulting logs, since variograms do not require the data to be evenly spaced. Due to the relatively large number of boreholes represented here, we assume that these boreholes provide a representative sample of the spatial statistics in Los Angeles basin.

Before performing the inversion on the data, we carry out a sensitivity test of the processing parameters. The left panel in Fig. 5 shows the variation in estimated parameters as a function of the window size of the median filter. The inset shows the smoothed log apal1 from Fig. 3 for different median filter dimensions plotted against the un-processed log (gray). The right panel shows the calculated variograms with different filter dimensions.

For the inversion, we first computed the expected variogram estimate for the Los Angeles basin from the geometric mean of the 38 individual variograms estimated from the sonic logs with the mean removed. After a suitable estimate for a combined Los Angeles variogram was obtained, we performed Monte Carlo simulations with 10,000 realizations of each combination of parameters for the von Karman autocovariance function. We then estimated variograms from the 10,000 realizations and computed the geometric mean to produce the expected synthetic variogram estimate. The synthetic variogram estimate was then compared with that from the borehole data. The best fitting model corresponds to a minimized chi-square GOF between the model and data.

The final inversion results for the synthetics and the modeled data are shown in Fig. 6. Our results show that a model with von Karman parameters ν=0.05 and a=50 m best replicate the data, while the ranges ν=0.0-0.1 and a=50-150 m are consistent with the data within the uncertainty. We directly estimate the standard deviation from the mean for each borehole log as ~5±2.5%. A comparison to the classical variogram estimate by Matheron (1963) shows that the smoother estimate using the moving-window variogram provides a more accurate input for our inversion of the data. Also, as a quality check on the analysis of the entire suite of logs, we estimate the variogram from the apal1 log using the Li and Lake (1994) method (Fig. 3c) assuming a fBm model. We estimate the Hurst exponent to be approximately 0.01, and the correlation length to be approximately 75 m, consistent with the results for the ensemble of sonic logs.

As the borehole data typically represents a suite of geological formations with different characteristics, we searched the data set for any depth-dependence of the Hurst exponent, correlation length, and standard deviation. Fig. 7 shows the moving-window variogram applied to the V_p ensemble data separated into 500 m depth bins. The results indicate no significant depth dependence in the variograms, and we conclude that treating the statistics with an anisotropic single-fractal model is sufficient to replicate the first-order effects of small-scale heterogeneities on the data.

GROUND MOTION SIMULATIONS

In the following sections we examine the scattering effects generated by statistical distributions of small-scale heterogeneities, and address the relative contribution of site effects and scattering along the path, scattering effects in the shallow sediments versus the entire model domain, and differences in the nature of scattering as a function of the depth of the source.

Modeling Small-Scale Heterogeneities

We generate the small-scale heterogeneities in the frequency domain using the following form for the amplitude portion of the power spectrum for 3D:

$$P(k) = P_0 (1 + 2\pi k a^2)^{-(1.5+v)} \quad , \tag{1.7}$$

where **k** is the magnitude of the wave number computed as $\mathbf{k} = \sqrt{k_x^2 + k_y^2 + k_z^2}$, and P_0 is a constant. We generate a normally distributed random field, and tune the amplitude spectrum according to Eq. 1.7, while leaving the phases random. This produces a power-law distribution of the power spectrum, as well as the time-domain autocorrelation function. We simulate anisotropy by stretching the velocity model with a horizontal-vertical anisotropy ratio and resample the resulting grid in the frequency domain to preserve the power spectrum of the original domain. This distribution can then be superimposed on existing 3D velocity models for southern California, which already contains longer-wavelength variation in the crust. We perturb both V_p and V_s (in terms of slowness, in order to preserve travel times) as well as densities by the same amount, in lack of more accurate constraints. However, tests indicate that the first-order effects on the ground motion are rather insensitive to these choices. Fig. 8 shows depth-sections (z = -160m), a cross-section, and 1D depth-profile from select locations in the Los Angeles basin for the SCEC CVM-SI 4.26 superimposed with a statistical distribution generated from the parameters constrained in our data analysis.

Comparing Synthetics and Data

In addition to a visual check of the similarities and differences between synthetics and data we use a series of GOF metrics to obtain an objective comparison. Here we use a subset of the metrics proposed by Olsen and Mayhew (2010), namely peak ground velocity (PGV), peak ground acceleration (PGA), and cumulative kinetic energy (CKE). Moreover, we calculate two additional GOF metrics Arias Intensity (AI, Arias, 1970) and the integral of the one-sided envelope of the seismogram. We calculate the AI as

$$I_A = \frac{\pi}{2g} \int_0^{T_d} \mathbf{a}(t)^2 \, dt \quad , \tag{1.8}$$

which has been found to provide a fairly reliable parameter for the shaking required to trigger landslides, Arias (1970). Here, T_d is the total duration of the seismogram. We chose these metrics as they are expected to be sensitive to the scattering effects from small-scale heterogeneities.

We compute GOF following Eq. 1.9

$$\text{GOF} = 100\,\text{erfc}[\text{NR}], \text{ where NR} = \frac{2|x-y|}{x+y}, \tag{1.9}$$

where x and y are selected metrics. To obtain a final GOF score for each model, we first compute the score for each component at each individual station. We average across the components to obtain a GOF score for one particular station followed by an average of all the stations to obtain a simulation average. In some cases, we use vector magnitude across all three components. Scores range from 0 to 100, where 100 indicates a perfect fit. Following Olsen and Mayhew (2010), as well as Anderson (2004), we use the following classification of the GOF values: 80-100 - excellent fit, 65-80 - very good fit, 45-65 - fair fit, and 35-45 - poor fit. We consider anything below 35 as a very poor fit.

Modeling Parameters for the 2008 M_w5.4 Chino Hills Event

In order to minimize the interference from finite fault effects on the desired scattering in the surrounding medium we choose a relatively small earthquake source, namely the 2008 M_w5.4 Chino Hills event. While the importance of scattering is expected to increase with frequency and distance from the source, we simulate the majority of our tests for frequencies up to 2.5 Hz in a 56 km by 40 km by 24 km model domain including part of the Los Angeles basin, San Gabriel basin and part of Chino basin, due to computational limitations (Fig. 9). We incorporate a finite-fault source model from Shao et al. (2012) consisting of 140 subfaults based on a kinematic source inversion. The statistical model is added to the CVM everywhere except within a spherical region around the source of radius 5 km in order to preserve the moment of the source for different small-scale perturbations in shear moduli. We use a constant grid spacing of 16m and a time step of 0.001s, for a total of 100 seconds of wave propagation. Each simulation used 13,125 processors on the Blue Waters Supercomputer at NCSA for a total of 6.5 wall clock hours (Table 1). We use AWP-ODC (Cui et al., 2010), a fourth-order staggered grid finite difference solution to the 3D elastic wave equation to simulate wave propagation. The code includes frequency-dependent attenuation (Withers et al., 2013) parameterized as $Q_s(f)=Qs_0f^{0.6}$, with Qs_0 constant below 1.0 Hz. We use the exponent n=0.6 for the power law as an average from the data constraints found by Song and Jordan (2013) for southern California. The representation of the earth's crust is provided by the SCEC UCVM package (Gill et al., 2014) using the CVM SI-4.26 velocity model (Lee et al., 2014). This particular model provides the most up-to-date representation of the Southern California crust including results from boreholes, 3D tomography, and full-waveform inversions.

Sensitivity Testing of the Statistical Models

In the following, we present the results of sensitivity tests on the ground motion synthetics for five parameters controlling the scattering from the small-scale heterogeneities: Hurst exponent, correlation length, standard deviation, seed number, and the horizontal-to-vertical anisotropy factor. The tests are conducted using the Chino Hills earthquake source, using $Qs = Qs_0f^{0.6}$ with $Qs_0=100V_s$ (V_s in km/s), followed by analysis

of the effects of varying Qs_0. We then examine the relative scattering contributions from the immediate area around the site, the sedimentary basin, and the underlying crust. Finally, we explore the scattering effects from the depth of the source and its proximity to a basin boundary.

Effects of Parameters of the Statistical Distributions

Fig. 10 shows horizontal slices of V_s in the 3D crustal model including the small-scale heterogeneities. Figs 11-12 show comparisons of 0-2.5 Hz waveforms and cumulative kinetic energies for six select stations in the model domain (LTP, RUS, SRN, STS, FUL and PSR). These stations sample different propagation distances and basin depths to provide a representative sample from the Los Angeles basin. We find that the Hurst exponent has little effect on the ground motions. Anisotropy of the small-scale heterogeneities has a small-to-moderate effect on the energy arriving at the stations, with higher energies for the larger anisotropy, because of the greater scattering in such model. Depending on the specific station, the correlation length can provide a greater variation of ground motions. We find lower peak ground motions for a = 5000 m as compared with a = 150 m which can be explained by the 3D velocity structure provided in the SCEC CVM-SI 4.26. In other words, the correlation length can be thought of the as the size of the heterogeneities. Here, the 5 km-scale heterogeneities are accounted for by the underlying velocity structure from the CVM, leaving the statistical distribution with a=150 m (as constrained by data, see section **Deep Borehole Measurements**) for the small-scale heterogeneities absent in the CVM.

The standard deviation parameter is a measure of the fluctuations of the heterogeneities from the mean background model, controlling the ground motion amplification along the wave propagation path. While the effect is relatively large, doubling the standard deviation from 5% to 10% does not necessarily cause peak motions and energy twice as large. In fact, in some cases (STS), the increase in standard deviation actually reduces the synthetics. As pointed out by Hartzell et al. (2010), this is due to a modified ray path through the medium, which can decrease the scattering. Finally, the seed number controls the spatial locations of the small-scale heterogeneities, creating a variance between the different realizations of the statistical model. The relatively large variation in ground motion energy at the 6 sites for different seed numbers (largest at FUL) implies that ensemble averages are needed for ground motion estimation when we incorporate statistical descriptions of the small-scale heterogeneities. We find that 5 or more models provide a satisfactory reduction in variance toward characterizing the expected ground motions through random scattering media.

Low-frequency Anelastic Attenuation

While more complicated representations of Q have been proposed (e.g., Taborda and Bielak, 2013) we aim at estimating the optimal linear long-period relationship between Qs_0 and V_s in this study. A widely used linear relation is Qs_0/V_s (km/s) = 50 (e.g., Graves and Pitarka, 2010). Here, we test the relations Qs_0/V_s = 50, 100, and 150, with a power law exponent of 0.6 for frequencies higher than 1 Hz (constrained by Song

and Jordan, 2013), and discuss the resulting GOF using 110 strong-motion stations. The simulations use the Chino Hills earthquake source in a model consisting of random perturbations from a von Karman model with parameters $\nu = 0.05$, $a = 150$ m, $\sigma = 5\%$, and H/V = 5 superimposed on SCEC CVM-SI 4.26. Figs. 13-15 show ensemble averages (5 realizations) of the simulations relative to data as a function of distance from the source as well as an interpolated map of the corresponding GOF for PGA, PGV, CKE and AI. The results for $Qs_0/V_s = 50$ (Fig. 13) have an average GOF score of 29 ('very-poor'). We see good fits in close proximity to the source (<15 km), with a rapid decay in GOF with further distance. $Qs_0/V_s = 100$ (Fig. 14) generates improved fits and distance trend, while still under-predicting the peak ground parameters. The model average GOF of 40 falls in the 'poor' category, with good fits at many stations, even at far distances from the source. The model using $Qs_0/V_s = 150$ (Fig. 15) has a better overall GOF score of 43 and good distance trend, but reduces some 'excellent' stations to 'average' and 'poor' scores close to the source. Fig. 16 compares synthetic with observed seismograms and cumulative kinetic energies at three representative basin sites for the three attenuation tests.

Figs. 13-16 show that our simulations under-predict the peak ground parameters (PGA, PGV, CKE, and AI). This can be attributed to an inadequate source description or material model to properly accommodate high-frequency seismic wave propagation. Our results suggest that the commonly used relation $Qs_0/V_s = 50$ may generate too strong attenuation for southern California, as least when considering the Chino Hills event. It is also clear that ensembles of statistical distributions of small-scale heterogeneities with parameters constrained from data can only account for a small fraction of the misfit between data and synthetics for the Chino Hills earthquake.

Site Versus Path Effects

It is well known that strong ground motions are strongly affected by the variation of near-surface low-velocity material (e.g., Anderson et al., 1996; Day, 1996; Boore and Joyner, 1997). For this reason we examine the contribution to the site effects from the statistical distributions of small-scale heterogeneities superimposed on the CVM-SI4.26 for the Chino Hills event. Particularly, we use simulations with the top 200 meters of the ground unperturbed by the statistical models and compare to results from simulations with heterogeneities added in the entire domain (Fig. 17). We see only small differences in the GOF estimates for the two models, with the model including small-scale heterogeneities ($\nu=0.05$, a=150m, H/V=5, $\sigma=5\%$, seed=1) throughout the entire domain providing slightly better fits on average. The small-scale heterogeneities in the upper 200 m account for amplifications and de-amplifications of up to 50%, as measured by the cumulative kinetic energy. The differences appear in very isolated and scattered locations, consistent with the observations of site effects from ground motion observations. However, the relatively small differences indicate that a majority of the scattering comes from the path of the seismic waves from the source location, as opposed to the immediate area around the site.

Depth-dependence of Small-scale Heterogeneities

We now examine the relative contribution of scattering from the near-surface low-velocity basin sediments using ensemble averages. Fig. 18 shows the results from simulations including statistical heterogeneities (v=0.05, a=150m, H/V=5, σ=5%, seed=1) added to the entire material model as well as only to sediments with V_s < 1500 m/s. For this particular model, there is very little difference in GOF between data and simulations. The model including heterogeneities in the entire basin (average GOF = 41) shows marginally improved fits across the entire model as compared to that for the model including small-scale heterogeneities for V_s<1500 m/s (average GOF = 39). However, the ratio between the cumulative kinetic energy for the two models varies by up to 50% (Fig. 18c). These results show that roughly equal contributions of the energy partitioned by crustal scattering originate in the lower velocity sediments and the underlying crust.

Depth-dependence of source

Olsen and Jacobsen (2011) explored the scattering from a shallow (0-5 km) SH-wave source horizontally and vertically incident onto a 1D basin structure with a statistical model of small-scale heterogeneities superimposed. The shallow horizontally incident source generated banded patterns of amplification radiating from the source into the basin, with amplification factors up to 4 compared to a reference model without the small-scale heterogeneities. While intriguing, these banded amplification patterns were generated using simulations with a somewhat unphysical earthquake source impinging on a simplified (square) basin model with the statistical model superimposed. Here, we further explore these results using a small shallow source on the boundary of the basin and surrounding rock. This model is meant to represent a small section of surface rupture on a range-bounding fault (e.g., San Andreas fault by the San Bernardino basin). Fig. 19 shows ratios of PGVs obtained from simulations with and without statistical distributions of small-scale heterogeneities generated by the parameters constrained by data, simulated in the same domain used for our Chino Hills modeling.

The results show that the bands of amplification are generated independently of the focal mechanism (including that for Chino Hills and pure strike-slip and dip-slip mechanisms) of a small near-surface (depth of 32 m) source located at the boundary of the basin sediments and rock, although the patterns of the bands change. However, when the shallow source is moved a few km away from the sediment-rock boundary, and for deeper sources, the bands of amplification deteriorate. This is likely caused by the breakup of the main wave front, before the interaction with the sediment-rock boundary. Our results imply that surface rupture on a range-bound fault (e.g., the San Andreas fault by the San Bernardino basin) may generate different patterns of ground motion shaking along lines parallel to the fault as compared to profiles perpendicular to the fault.

Additional Causes of GOF Misfit for Chino Hills

The tests described above show that while the statistical models of small-scale heterogeneities can generate localized strong amplifications and de-amplifications of the ground motions, they cannot explain the discrepancies to data for the Chino Hills event, as measured by PGVs, PGAs, CKEs and AI. In the following, we perform two additional tests, without including statistical models, to try to explain the discrepancies.

Effects of a V_s30-derived GTL

Ely (2010) proposed to add a GTL, derived from V_s30 measurements (Wills and Clahan, 2006) to the SCEC CVMs, in order to make up for insufficient resolution in the near-surface regions of the velocity models. Figure 20 shows comparisons of PGVs, PGAs, CKE and AI for results with and without the GTL. In general, the GOF values tend to increase (up to 20% for PGVs and PGAs, and up to 50% for CKE and AI) when the GTL layer is included. The average GOF for the metrics is increased by 3% when the GTL is included. No statistical models are included in these results.

Source Magnitude

The finite fault description of the Chino Hills event from Shao et al. (2012) corresponds to a moment release of $1.58 \ 10^{17}$ Nm, in agreement with the USGS Centroid estimate and corresponding to $M_w5.4$. However, considering that estimates of source moments can vary by up to a factor of two, we have explored how moment increases by 25% ($M_w5.50$) and 50% ($M_w5.55$) relative to that provided by Shao et al. (2012) affect the GOF for Chino Hills (Fig. 21). It is clear that an increase in the moment of at least 25% but less than 50% increases the GOF by an amount much larger than was the case for the statistical models of small-scale heterogeneities.

Conclusions

We have used V_s30 data measurements and borehole sonic log data to constrain the parameters describing statistical distributions of small-scale heterogeneities in the Los Angeles basin. Our results show that the von Karman parameters most representative of the velocity variability are $\nu = 0.05 - 0.1$, $a = 50 - 150$ m, $H/V = 5$, and $\sigma = 5 \pm 2.5\%$. The parameters ν and a are primarily constrained by inverting 38 borehole logs located around the Los Angeles area, as the V_s30 data provide limited resolution. We estimate σ by directly computing the standard deviation of the velocity with respect to the mean of the each borehole, and averaging over the entire basin. We lack the data resolution to properly constrain the horizontal-to-vertical anisotropy (H/V) but our findings are consistent with H/V = 5 based on previous studies (e.g., Holliger and Levander, 1992) and ground motion validation of the statistical models using wave propagation simulations. Moreover, these statistical parameters constitute an accurate data-driven

model that can be used for high-frequency simulations of seismic wave propagation in Southern California.

Next, we simulate 3D linear viscoelastic waves through random media with frequency-dependent attenuation using a 4th order staggered-grid finite-difference solution to the elastodynamic wave equation. We compare synthetic with observed seismograms at strong-motion stations in the Los Angeles basin, to gauge our understanding of wave propagation through realistic 3D velocity structures including small-scale heterogeneities. We find that the Hurst exponent has little effect on the ground motions. Increases in correlation length and horizontal-vertical anisotropy (from 2 to 5) of the small-scale heterogeneities produce more scattering in the model, but do not necessarily result in more energy arriving at the station due to the complex nature of path effects. The standard deviation, constrained by the data to be 5% throughout the entire basin, has a large effect on the resulting ground motions. Finally, the seed number introduces an inherent variance on the resulting ground motions, which is caused by the spatial relocation of the velocity perturbations. Due to the random nature of the fractal distribution, care must be taken when comparing synthetics with observed seismograms. We recommend at least five simulations per ensemble to reduce the variance associated with differences due to random realizations of the statistical distribution.

We find when including low shear-wave velocities (V_s<200 m/s) coupled with contributions from small-scale heterogeneities changes that the commonly used relation Qs_0/V_s (km/s)=50 for southern California (Graves and Pitarka, 2010) may provide too strong attenuation of the waves. Based on select metrics (PGA, PGV, CKE and AI versus distance against strong motion data, we find improved GOF scores against strong motion data for Qs_0/V_s (km/s) =100-150.

In an effort to understand the contributions from the small-scale heterogeneities, we perform tests to address the following questions: (1) what is the relative contribution from site and path scattering effects, (2) is there a depth-dependence of the statistical model – does the majority of scattering occur near the earth's surface or at source depths, and (3) what are the effects of the source location on amplifications/de-amplifications from crustal scattering? We find that the small-scale heterogeneities in the upper 200 m account for amplifications and de-amplifications of up to a factor of two, as measured by the cumulative kinetic energy. The differences appear in very isolated and scattered locations, consistent with the observations of site effects from ground motion observations. Furthermore, we find approximately equal contributions to the scattering from the basins and the underlying crust. Thus, small-scale heterogeneities must be accounted for throughout the entire modeling domain.

Finally, we find unique banded amplification PGV patterns generated by small-scale heterogeneities when we introduce a source near the surface (z = -32m) at a basin/bedrock interface. The results show that the bands of amplification are generated independently of the focal mechanism (including that for Chino Hills, and pure strike-slip and dip-slip sources) although some characteristics of the bands change. Our results imply that surface rupture on a range-bound fault (e.g., the San Andreas fault by the San

Bernardino basin) may generate different patterns of ground motion shaking along lines parallel to the fault as compared to profiles perpendicular to the fault.

Our sensitivity simulations and validation for the Chino Hills event have revealed many details on shallow crustal scattering phenomena as well as constraints on parameters to model the resulting ground motions using statistical distributions. The small-scale scattering clearly cannot account for the misfit between data and synthetics for the Chino Hills event, which is more likely due to other deficiencies in the crustal model (e.g., lack of a GTL), and/or the source description (e.g., underestimate of moment, energy of higher frequencies). In any case, we expect that crustal scattering due to small-scale heterogeneities will play an increasingly important role in ground motion prediction as regional-scale simulations, at higher frequencies, become commonplace for physics-based seismic hazard analysis.

Acknowledgements

References

Anderson, J. G. (2004). Quantitative measure of the goodness-of-fit of synthetic seismograms. *Proc. 13th World Conf. on Earthquake Eng., Vancouver, British Columbia, Canada, Int. Assoc. Earthquake Eng.*, Paper 243.

Anderson, J. G., Y. Lee, Y. Zeng, and S.M. Day (1996). Control of strong motion by the upper 30 meters. *Bulletin of the Seismological Society of America*, **86(6),** 1749–1759.

Arias, A. (1970). Measure of earthquake intensity. Massachusetts Inst. of Tech., Cambridge. Univ. of Chile, Santiago de Chile.

Boore, D. M., and W.B. Joyner (1997). Site amplifications for generic rock sites. *Bulletin of the Seismological Society of America*, **87(2),** 327–341.

Cui, Y., K.B. Olsen, T. H. Jordan, K. Lee, J. Zhou, P. Small, D. Roten, G. Ely, D.K. Panda, A. Chourasia, J. Levesque, S. M. Day, and P. Maechling (2010). Scalable earthquake simulation on petascale supercomputers. In *High Performance Computing, Networking, Storage and Analysis (SC), 2010 International Conference for* (pp. 1-20). IEEE.

Day, S. M. (1996). RMS response of a one-dimensional half-space to SH. *Bulletin of the Seismological Society of Bulletin of the Seismological Society of America*, **86(2),** 363–370.

De, G. S., D. F. Winterstein, and M.A. Meadows (1994). Comparison of P- and S-wave velocities and Q's from VSP and sonic log data. *Geophysics*, **59(10),** 1512–1529. doi:10.1190/1.1443541

Dolan, S., C. Bean, and B. Riollet (1998). The broad-band fractal nature of heterogeneity in the upper crust from petrophysical logs. *Geophysical Journal International*,

132(3), 489–507. doi:10.1046/j.1365-246X.1998.00410.x

Ely, G., T. H. Jordan, P. Small, P. J. Maechling (2010). A V_s30-derived Near-surface Seismic Velocity Model Abstract S51A-1907, presented at 2010 Fall Meeting, AGU, San Francisco, Calif., 13-17 Dec.

Frankel, A., and R. Clayton (1986). Finite difference simulations of seismic scattering: Implications for the propagation of short-period seismic waves in the crust and models of crustal heterogeneity. *Journal of Geophysical Research: Solid Earth Geophysics*, **91(B6)**, 6465–6489.

Gill, D., P. Small, P. Maechling, T.H. Jordan, A. Plesch, R. Taborda, and S. Callaghan (2014). UCVM: Open Source Software Framework for 3D Velocity Models, *Seism. Res. Lett.* **85:2**, 457.

Graves, R.W., and A. Pitarka (2010). Broadband ground-motion simulation using a hybrid approach, *Bull. Seis. Soc. Am.* **100**, 5A, 2095-2123, doi: 10.1785/0120100057.

Hartzell, S., S. Harmsen, and A. Frankel (2010). Effects of 3D Random Correlated Velocity Perturbations on Predicted Ground Motions. *Bulletin of the Seismological Society of America*, **100(4)**, 1415–1426. doi:10.1785/0120090060.

Holliger, K. (1996). Upper-crustal seismic velocity heterogeneity as derived from a variety of P-wave sonic logs, *J. Geophys. Int*., **125**, 813-829.

Holliger, K. (1997). Seismic scattering in the upper crystalline crust based on evidence from sonic logs. *Geophysical Journal International*, **128(1),** 65–72.

Holliger, K., and A.R. Levander (1992). A Stochastic View of Lower Crustal Fabric Based on Evidence From the Ivrea Zone. *Geophysical Research Letters*, **19(11),** 1153–1156.

Imperatori, W., & Mai, P. M. (2013). Broad-band near-field ground motion simulations in 3-dimensional scattering media. *Geophysical Journal International*, **192(2),** 725–744. doi:10.1093/gji/ggs041.

Kohler, M. D., H. Magistrale, and R.W. Clayton (2003). Mantle Heterogeneities and the SCEC Reference Three-Dimensional Seismic Velocity Model Version 3. *Bulletin of the Seismological Society of America*, **93(2),** 757–774. doi:10.1785/0120020017.

Lee, E-J., P. Chen, and T.H. Jordan (2014). Testing waveform predictions of 3D velocity models against 2 recent Los Angeles earthquakes, *Seism. Res. Lett.,* in press (Nov/Dec 2014 issue).

Levander, A., R. England, S. Smith, R. Hobbs, J. Goff, and K. Holliger (1994). Stochastic characterization and seismic response of upper and middle crustal rocks based on the Lewisian gneiss complex, Scotland, *J. Geophys. Int.* **119**, 243-259.

Li, D., and L.W. Lake (2010). A moving window semivariance estimator. *Water Resources Research*, *30*(5), 1479–1489. doi:10.1029/94WR00319

Louie, J. (2005). Improving next-generation attenuation models with shear-velocity measurements at all TriNet and strong-motion station in LA. Final technical report submitted to NEHRP, U.S. Geological Survey.

Louie, J. (2007). Shear-wave velocity map for California: Collaborative research with CGS and UNR. Final technical report submitted to NEHRP, U.S. Geological Survey.

Magistrale, H., R. Clayton, and R.W. Graves (2000). The SCEC Southern California Reference Three-Dimensional Seismic Velocity Model Version 2. *Bulletin of the Seismological Society of America* **90(6B)**), S65-S76.

Matheron, G. (1963). Principles of geostatistics. *Economic geology*, **58(8),** 1246-1266.

Mela, K., and J.N. Louie (2001). Correlation length and fractal dimension interpretation from seismic data using variograms and power spectra. *Geophysics*, **66(5),** 1372–1378. doi:10.1190/1.1487083.

Olsen, K. B., and J.E. Mayhew (2010). Goodness-of-fit criteria for broadband synthetic seismograms, with application to the 2008 Mw 5.4 Chino Hills, California, earthquake. *Seismological Research Letters*, **81(5),** 715–723, doi:10.1785/gssrl.81.5.715

Olsen, K. B., S.M. Day, and C.R. Bradley (2003). Estimation of Q for Long-Period (>2 sec) Waves in the Los Angeles Basin. *Bulletin of the Seismological Society of America*, **93(2),** 627–638. doi:10.1785/0120020135

Olsen, K.B. and B.H. Jacobsen (2011). Spatial variability of ground motion amplification from low-velocity sediments including fractal inhomogeneities with special reference to the southern California basins, *in* Procs. 4th IASPEI/IAEE International Symposium *Effects of Surface Geology on Seismic Motions (ESG4),* Aug 23-26 2001, UC Santa Barbara.

Pullammanappallil, S., A. Levander, and S.P. Larkin (1997). Estimation of stochastic crustal parameters from seismic exploration data, *J. Geophys. Res.,* **102,** 15,269-15,286.

Shao, G., C. Ji, and E. Hauksson (2012). Rupture process and energy budget of the 29 July 2008 Mw 5.4 Chino Hills, California, earthquake, *Journal of Geophysical Research*. **117**, B07307, doi:10.1029/2011JB008856.

Song, X., and T.H. Jordan (2013). Anelastic attenuation and elastic scattering of seismic waves in the Los Angeles region, Procs. Southern California Earthquake Center annual meeting, Palm Springs, September 2013.

Süss, M. P., and J.H. Shaw (2003). P wave seismic velocity structure derived from sonic logs and industry reflection data in the Los Angeles basin, California. *Journal of Geophysical Research*, **108(B3),** 2170. doi:10.1029/2001JB001628

Taborda, R. and J. Bielak (2013). Ground-Motion Simulation and Validation of the 2008 Chino Hills, California, Earthquake. *Bulletin of the Seismological Society of America*, **103(1),** 131–156. doi:10.1785/0120110325.

Tatarski, V.I. (1961). Wave propagation in a turbulent medium, McGraw-Hill, New York.

Wills, C.J., and K.B. Clahan (2006). Developing a map of geologically defined site-condition categories for California, *Bull. Seis. Soc. Am.* **96(4A),** 148301501, doi:10.1785/0120050179.

Withers, K., K.B. Olsen, and S.M. Day (2013). Deterministic high-frequency ground motion using dynamic rupture along rough faults, small-scale heterogeneities, and frequency-dependent attenuation, Procs of the American Geophysical Union Fall Mtg, San Francisco, December 2013 (S52C-08).

Wu, R. S., Z., Xu, and X.-P. Li (1994). Heterogeneity spectrum and scale-anisotropy in the upper crust revealed by the German Continental Deep-Drilling (KTB) Holes. *Geophysical Research Letters,* **21(10),** 911–914. doi:10.1029/94GL00772

Yong, A., Martin, A., Stokoe, K., and Diehl, J. (2013). ARRA-funded V_s30
measurements using multi-technique approach at strong-motion stations in
California and central-eastern United States: U.S. Geological Survey Open-File
Report 2013–1102, 60 p. and data files,http://pubs.usgs.gov/of/2013/1102/.

Figure Captions

Figure 1. Map showing the 639 V_s30 measurements collected for use in the geostatistical
analysis. (Red) 350 measurements incorporated in the SCEC CVM-S 4.0 (Magistrale et
al., 2000), (blue) 102 measurements from Louie (2005, 2007), and (green) 187
measurements provided by Alan Yong (Yong et al., 2013).

Figure 2. Variogram estimated for the V_s30 measurements in the Los Angeles basin.
The dashed green line shows the logarithmic regression used to estimate the Hurst
exponent ($\nu = 0.33$, corresponding to a fractal dimension of $D = 1.67$). Correlation
lengths are estimated between 5 and 10 km.

Figure 3 a) V_p sonic log (gray) with long-period 100 m median filter (black), and 5 m-
smoothed log (red). b) De-trended fluctuation profile (blue) generated by subtracting the
long-period median filter (black) from the red and gray logs in a), and used to estimate
variograms. c) Variogram estimate from the fluctuation profile (blue) of the apal1
borehole shown in b). The fluctuation profile (blue) becomes shorter due to edge effects
from the median filter. The dashed green line is the logarithmic regression used to
estimate the Hurst exponent. We estimate the correlation length to be between 50 and 80
m, and $\nu=0.01$.

Figure 4. Map showing 38 deep borehole locations within the Los Angeles basin.

Figure 5. (Left) Sensitivity of the median filter applied to the sonic logs (gray, before
filtering, and red, after filtering) on the estimates of the von Karman parameters. As the
filter dimension increases, the resulting sonic log becomes smoother, increasing the Hurst
exponent estimates while decreasing the estimate of the standard deviation. (Right)
Variograms plotted for different filter sizes. Note that the filtering has little effect on
estimates of the correlation lengths.

Figure 6. Inversion results for best fitting von Karman parameters. Dashed lines show
variogram estimates using the classical estimate of Matheron (1963). The inset shows one
realization from the best fitting model trend plotted against the borehole data.

Figure 7. Depth-dependent analysis for the 38 analyzed V_p boreholes in the Los Angeles
basin. We see little variation in the statistical parameters at different depth intervals
indicating there is no significant depth dependence of the von Karman parameters.

Figure 8. Random perturbations from the von Karman model with parameters $\nu = 0.05$, a
$= 50$ m, $\sigma = 5\%$, H/V $= 5$ superimposed on SCEC CVM-SI 4.26. (a) Depth slice at 160 m
depth, (b) cross-section taken at A-A', and (c) profile extracted at the red circle in (a).
Velocities are listed in (m/s). Note in b) that the perturbations are excluded from the

source region to preserve the moment released during the earthquake. The star depicts the hypocenter for the Chino Hills event.

Figure 9. Topographic map showing the focal mechanism and simulation domain (rectangle, 56 km x 40 km x 24 km) for the 2008 M_w5.4 Chino Hills earthquake. Gray lines depict major freeways, and red lines are faults in the greater Los Angeles area.

Figure 10. Depth slices taken at z = -160m in the CVM-SI 4.26 superimposed with various statistical models. We vary the von Karman parameters across the defined parameter space from previous studies. Contour lines show V_s = 2500 m/s isosurface plotted in 2 km intervals.

Figure 11. Velocity waveforms and cumulative kinetic energy for synthetics and data at stations seismograms compared with data record at three stations (left) RUS, (center) LTP, and right (SRN). The following models from Fig. 10 were used: top row (b) and (c), second row (a) and (d), third row (f) and (a), fourth row (a) and (e), and bottom row (a) and (g).

Figure 12. Same as Fig. 11, but for stations (left) STS, (center) PSR, and (right) FUL.

Figure 13. Test of the influence of reference Q_s for the GOF. Comparison of (a) PGV, (b) PGA, (c) CKE, and (d) AI versus distance for ensemble averages (blue lines) from simulations of the M_w5.4 Chino Hills earthquake using anelastic attenuation defined as Qs_0/V_s=50 (V_s in km/s) and Qp_0=2Qs_0 to data (gray lines, with average in black). The simulations were carried out with a von Karman distribution of the small-scale heterogeneities with ν = 0.05, a = 150 m, σ = 5%, H/V = 5 superimposed on SCEC CVM-SI 4.26. The shaded regions show the variance of the different seed numbers. e) GOF score averaged for all four metrics. We see decent fits close to the source, with very poor fits at further distances from the source.

Figure 14. Same as Fig. 13, but for anelastic attenuation defined as Qs_0/V_s=100. All metrics are under-predicted at distances greater than about 10 km from the source. However, we see generally better fits at farther distances, as compared to the results for Qs_0/V_s=50 (Fig. 13).

Figure 15. Same as Fig. 13, but for anelastic attenuation defined as Qs_0/V_s=150. The fit is improved as compared to the results for Qs_0/V_s=50 and 100, in particular at distances from the source larger than 15 km (Figs. 13-14).

Figure 16. Influence of reference Q_s on seismograms and cumulative kinetic energy for (a) deep basin site LTP, (b) shallow basin site RUS, and (c) hard rock site SRN for simulations shown in Figs. 13-15.

Figure 17. GOF (CKE) for the Chino Hills source in the CVM-SI4.26 including small-scale heterogeneities (ν=0.05, a=150m, H/V=5, σ=5%, seed=1). (a) Small-scale heterogeneities added to where V_s < 1500 m/s, and b) in the entire model domain. (c) Ratios of CKE between a) and b), generating up to 50% variation between the simulations. (d) Comparison between cumulative kinetic energy for data (gray, smoothed in black), simulations with no heterogeneities added (green), simulations with the

heterogeneities added where V_s < 1500 m/s (blue), and simulations with heterogeneities added everywhere. The star depicts the epicenter of the Chino Hills event.

Figure 18. GOF (CKE) for the Chino Hills source in the CVM-SI4.26 including small-scale heterogeneities (ν=0.05, a=150m, H/V=5, σ=5%, seed=1). (a) Small-scale heterogeneities added everywhere except the top 200 m, and b) in the entire model domain. (c) Ratios of CKE between a) and b). (d) Comparison between CKE for simulations with no heterogeneities added (green), simulations with the heterogeneities added where V_s < 1500 m/s (blue), and simulations with heterogeneities added everywhere. The star depicts the epicenter of the Chino Hills event. (d) Comparison between cumulative kinetic energy for data (gray, and smoothed in black), simulations with no heterogeneities added (green), simulations with the top 200 m unperturbed (blue), and simulations with heterogeneities added everywhere (red).

Figure 19. Maps of PGV ratios between results from simulations with and without statistical models added using hypothetical sources in the Chino Hills model domain, to explore the effects of small-scale heterogeneities on near-surface ruptures. a) Pure horizontal strike-slip source at z=-16m directly on the interface between the basin and bedrock. b) Pure vertical dip-slip source at z=-16m directly. c) Chino Hills source mechanism in the same location as for a) and b). Panel d) pure strike-slip shear source located 4 km east of the rock-sediment interface. Notice the reduced banding caused by increasing the distance of the source to the rock-sediment interface.

Figure 20. Effects of adding a V_s30-based GTL to the CVM SI4.26. Comparisons are shown for (a) PGV, (b) PGAs, (c) CKE, and (d) AI. Average GOF maps for the four metrics in a-d are shown in for the simulations (e) without and (f) with the GTL. No statistical model is included in these simulations.

Figure 21. Effects of scaling the moment of the Chino Hills event, relative to that from Shao et al. (2012) of $1.58 \ 10^{17}$ Nm. Comparisons of 0-1 Hz synthetics and data are shown.

Tables:

Table 1. Simulation Parameters

Domain	
Length	56 km (3500 nodes)
Width	40 km (2500 nodes)
Depth	24 km (1500 nodes)
Southwest Corner	33.7500, -118.3000

Spatial Resolution	
Maximum Frequency	2.5 Hz
Minimum V_s	200 m/s
Points per wavelength	5
Grid Discretization	16 m
Number of cells	13.125 billion
Wall-clock time	6.5 hours
Number of Processors	13,125

Temporal Resolution	
Time Discretization	0.001 s
Simulation Time	100.0 s
Number of Timesteps	100,000
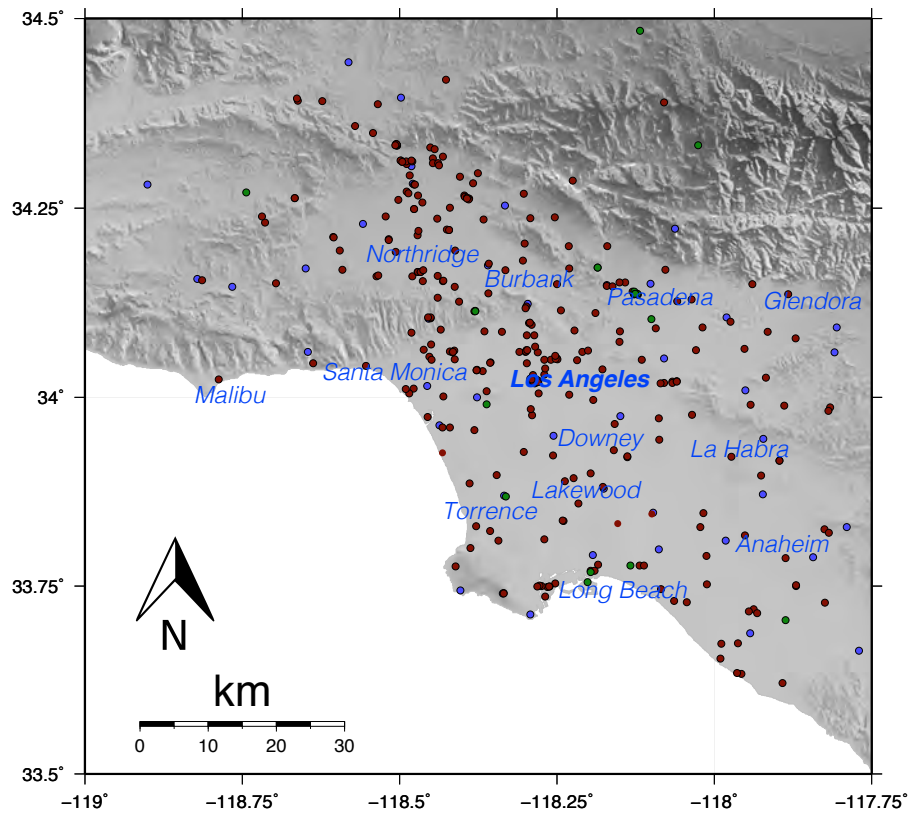
Figure 1.
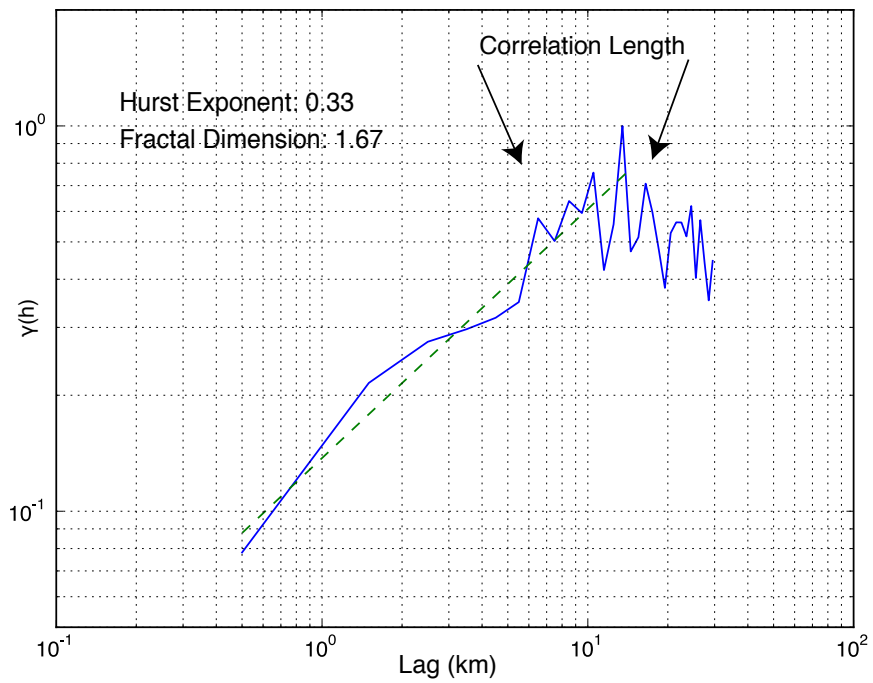
Figure 2.
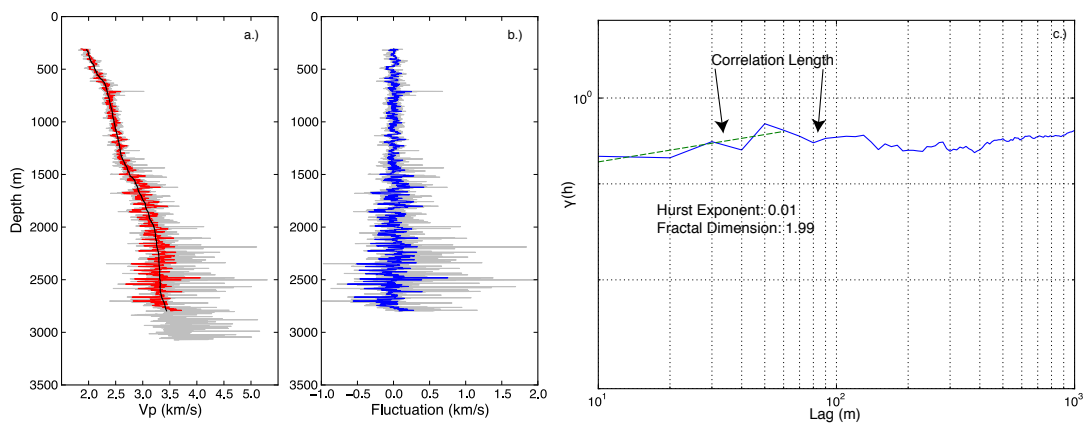
Figure 3.
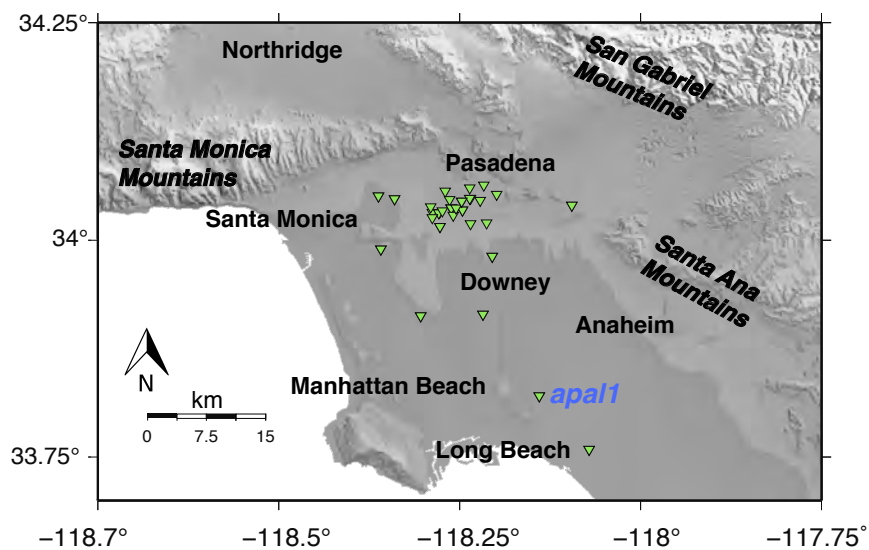
Figure 4.
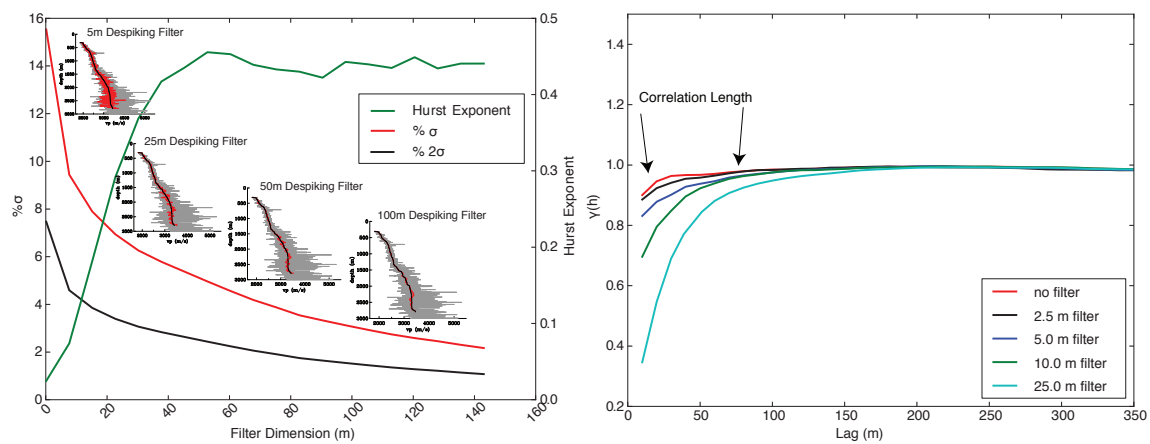
Figure 5.
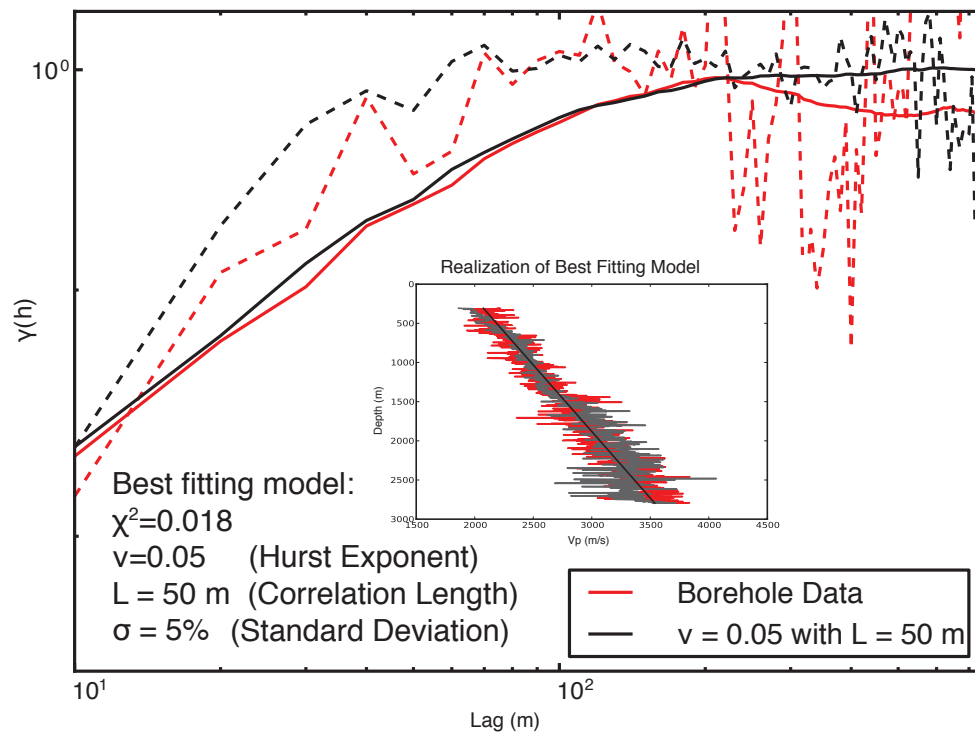
Figure 6.
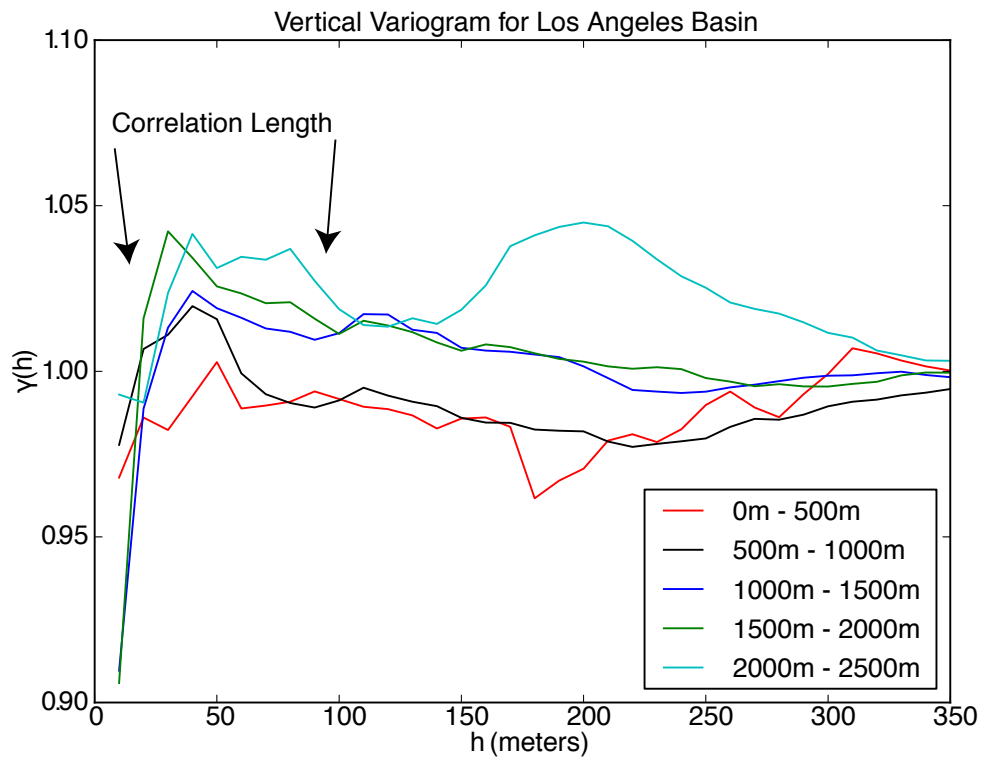
Figure 7.
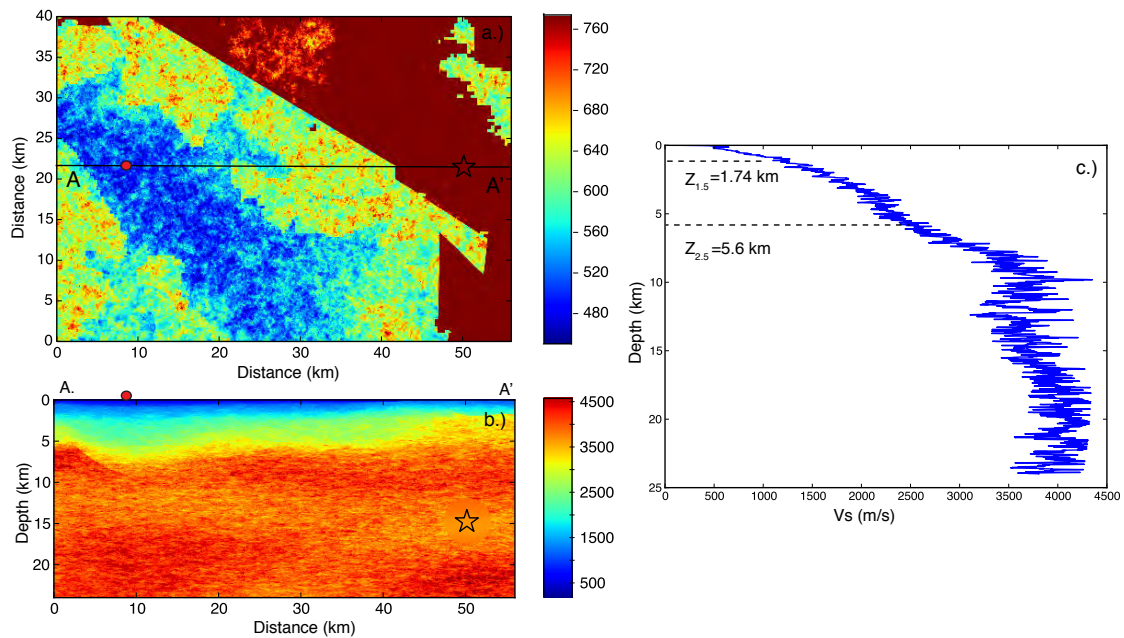
Figure 8.
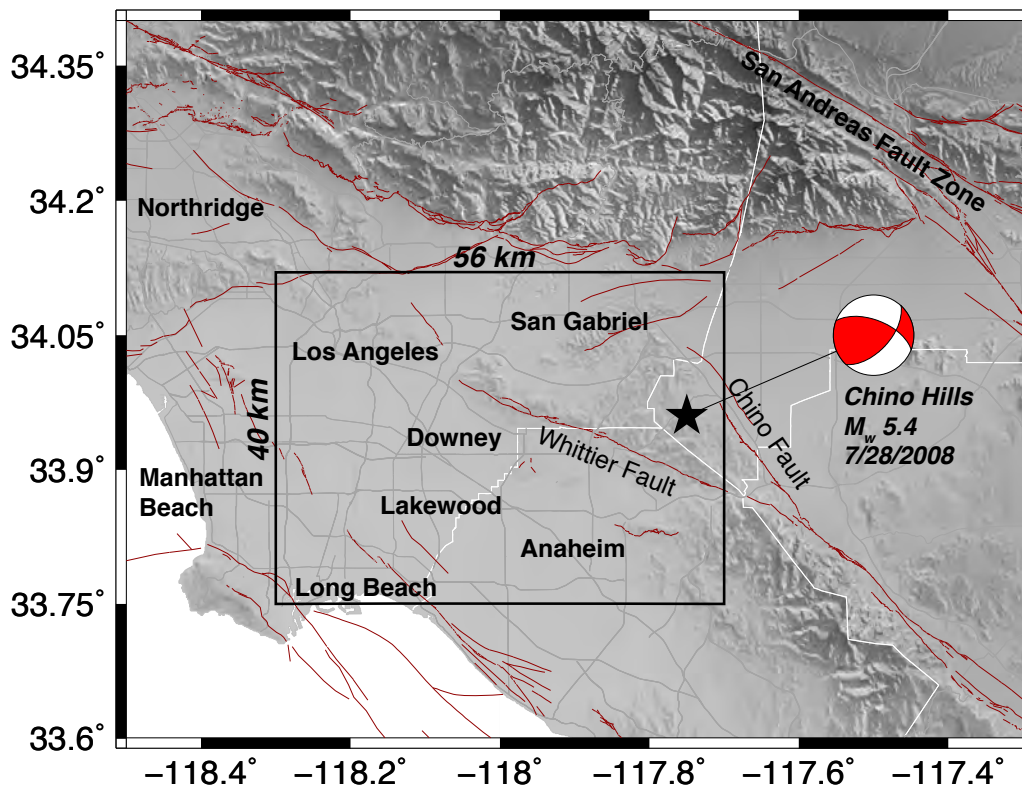
Figure 9.
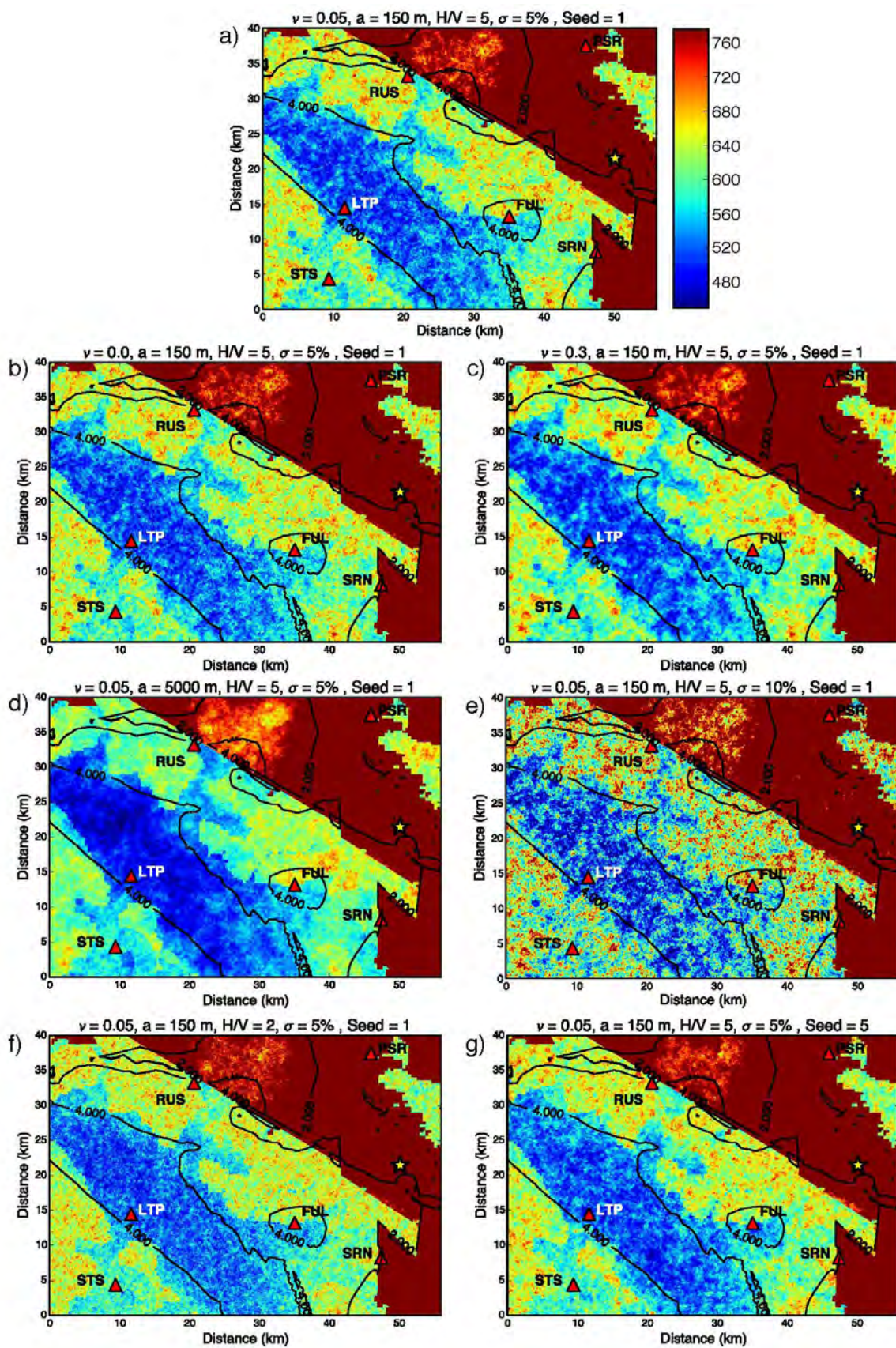
Figure 10.
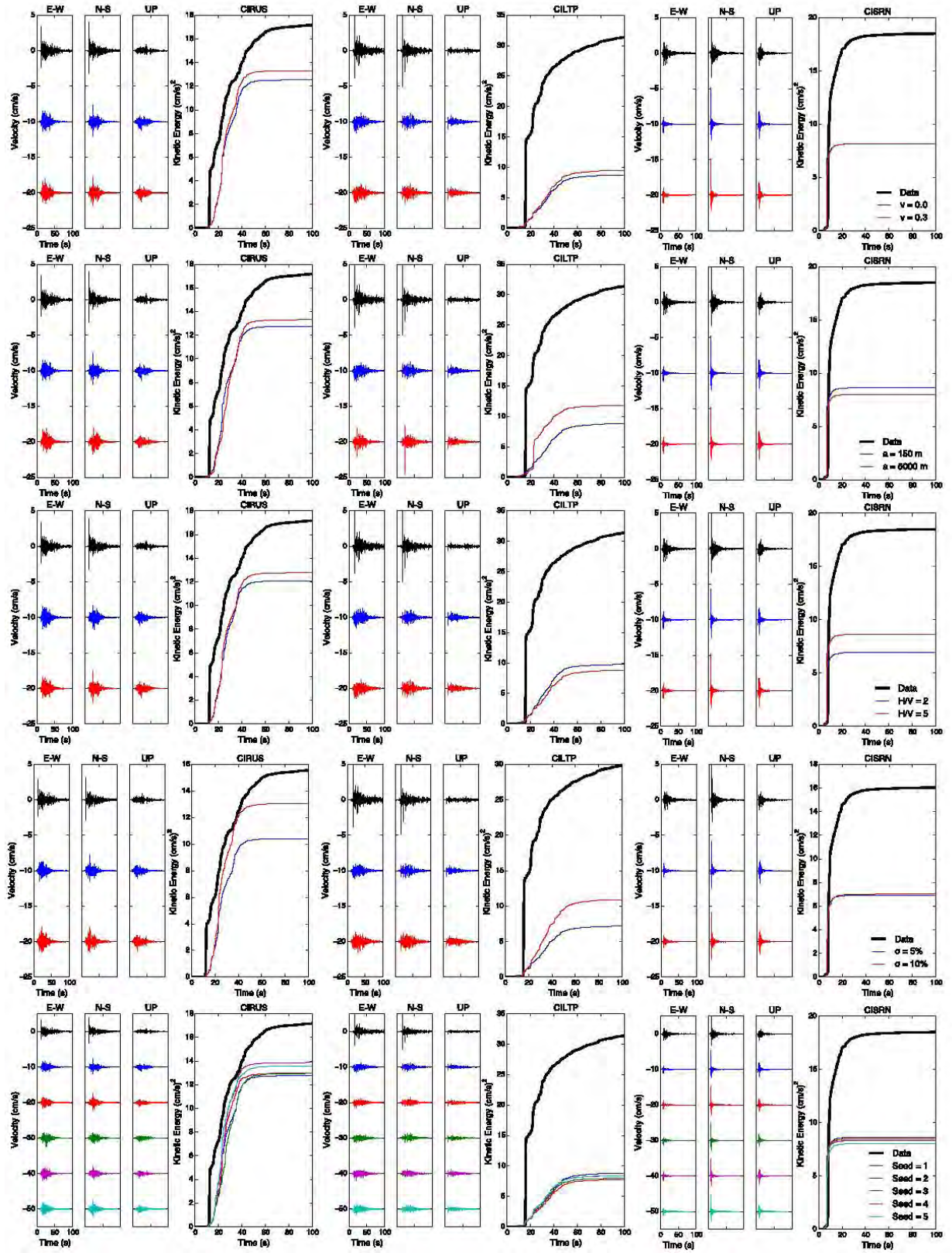
Figure 11.
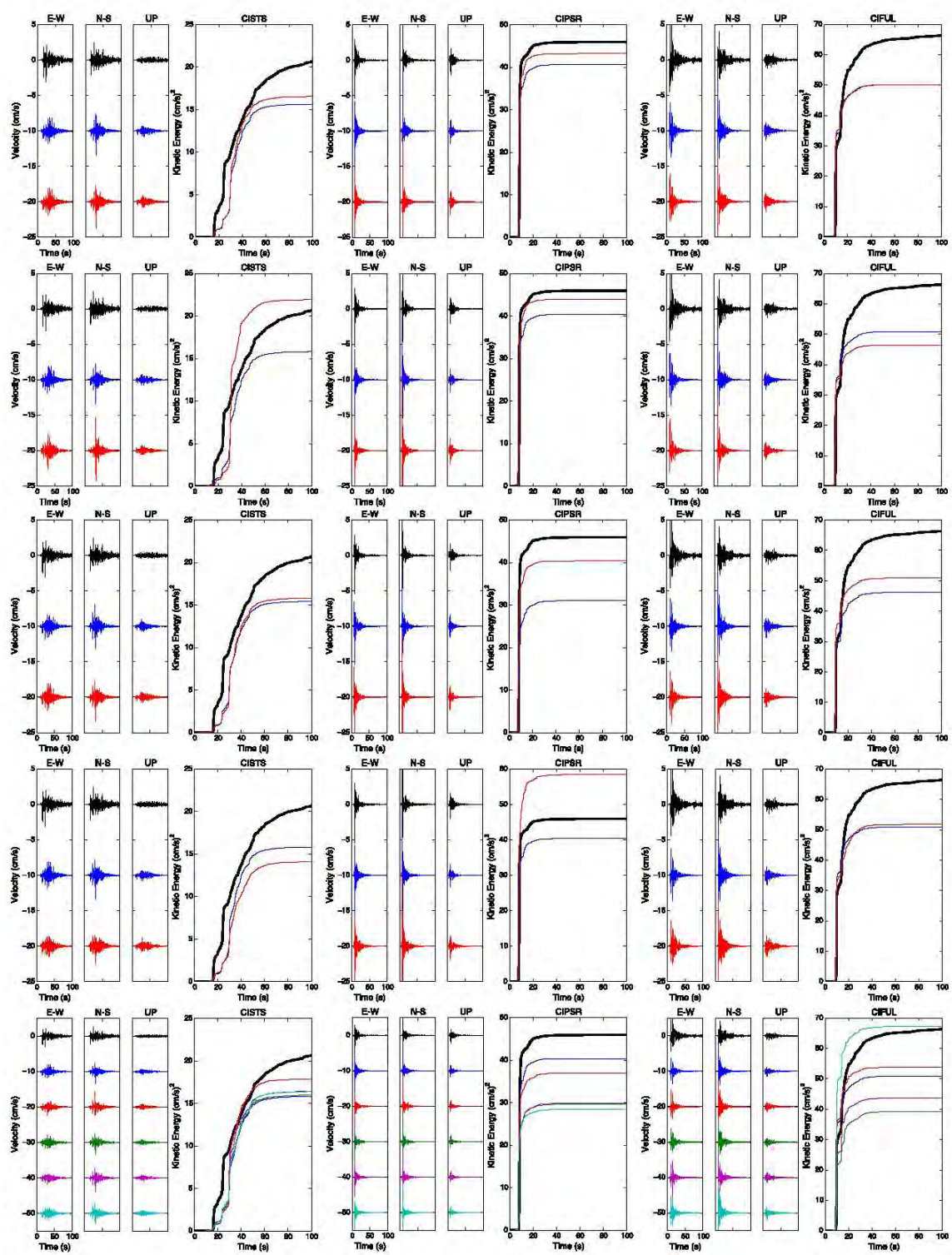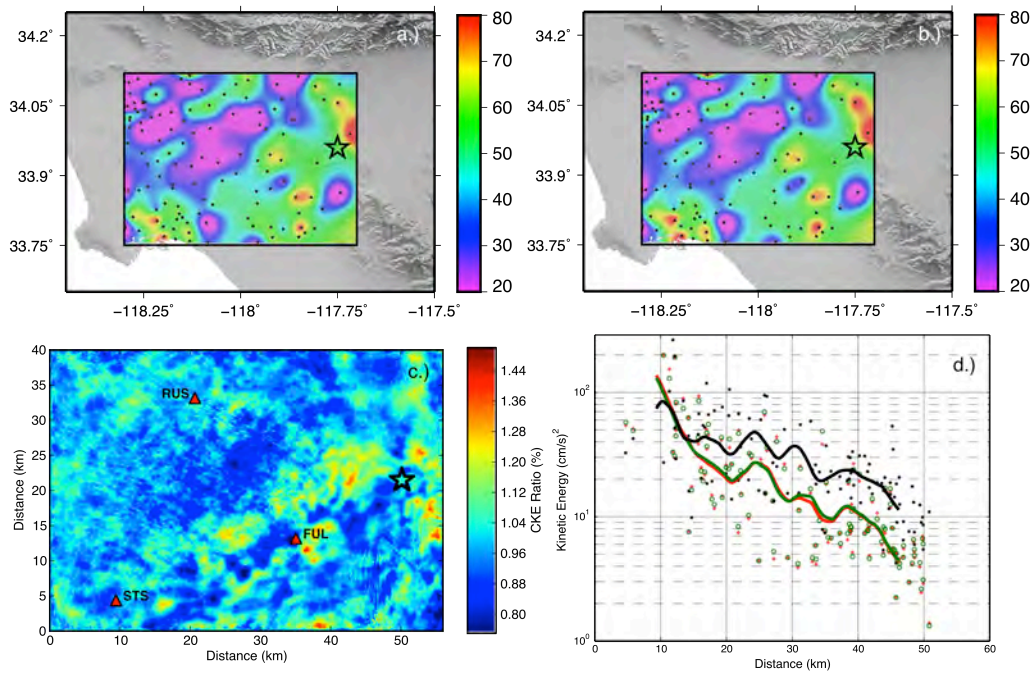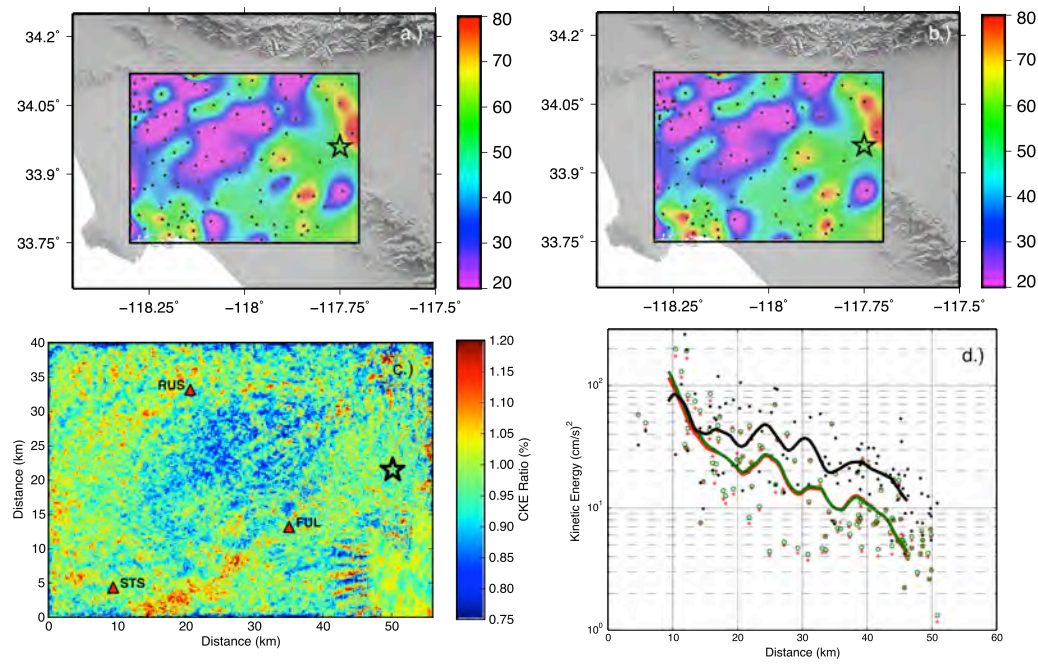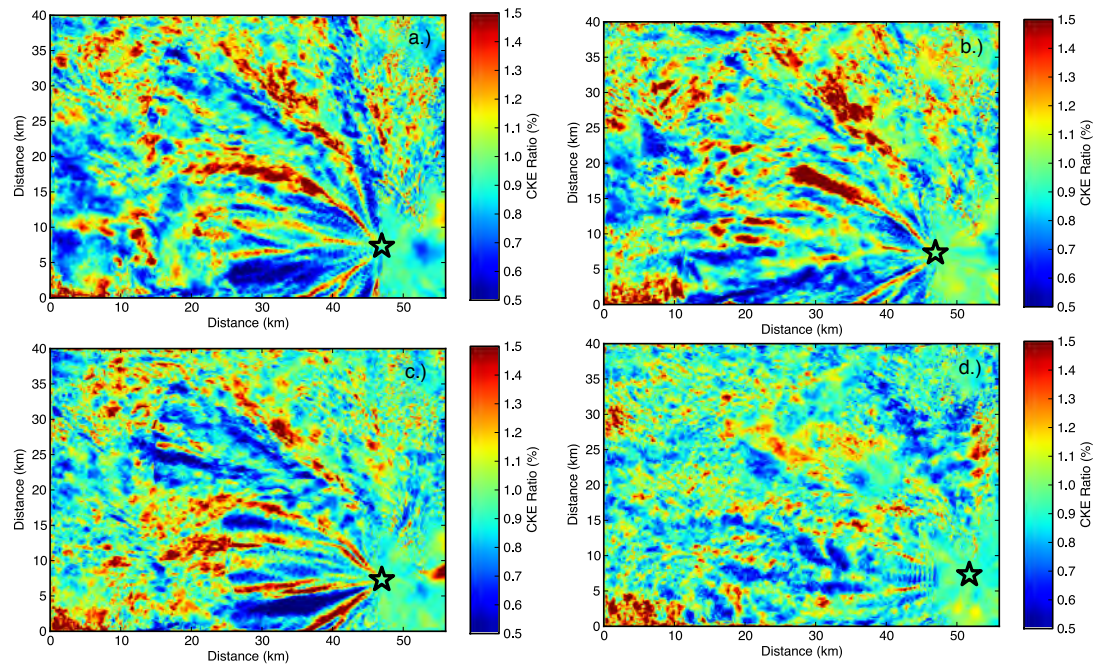
Figure 12.
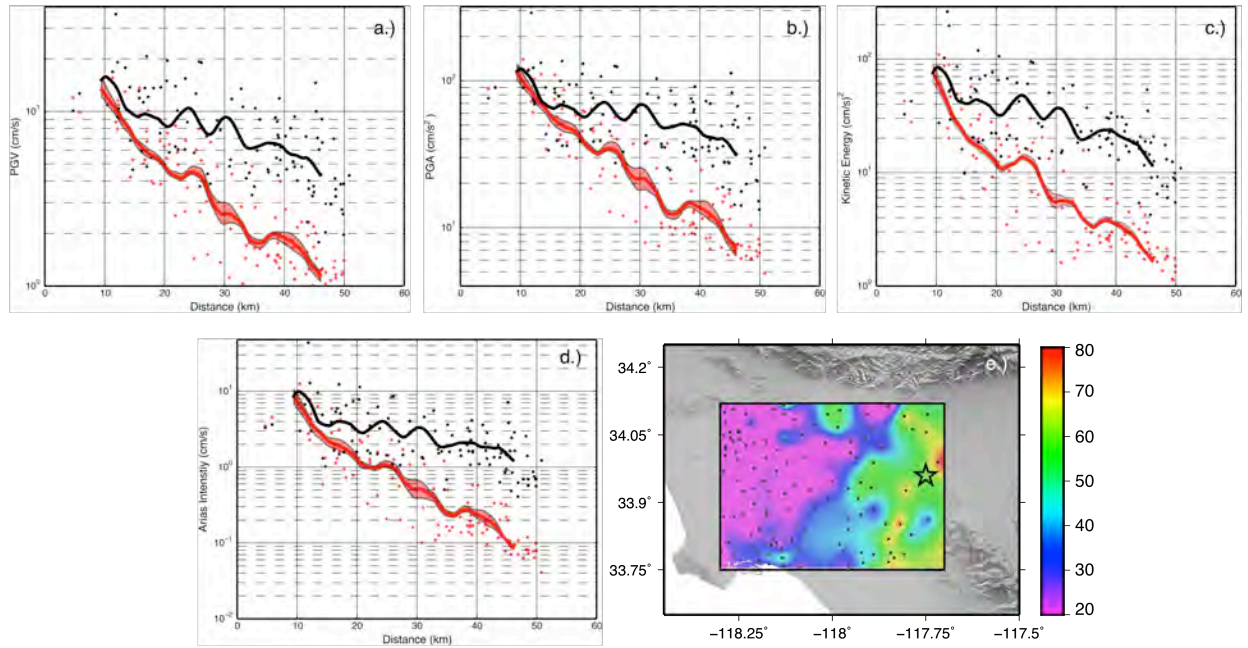
Figure 13.
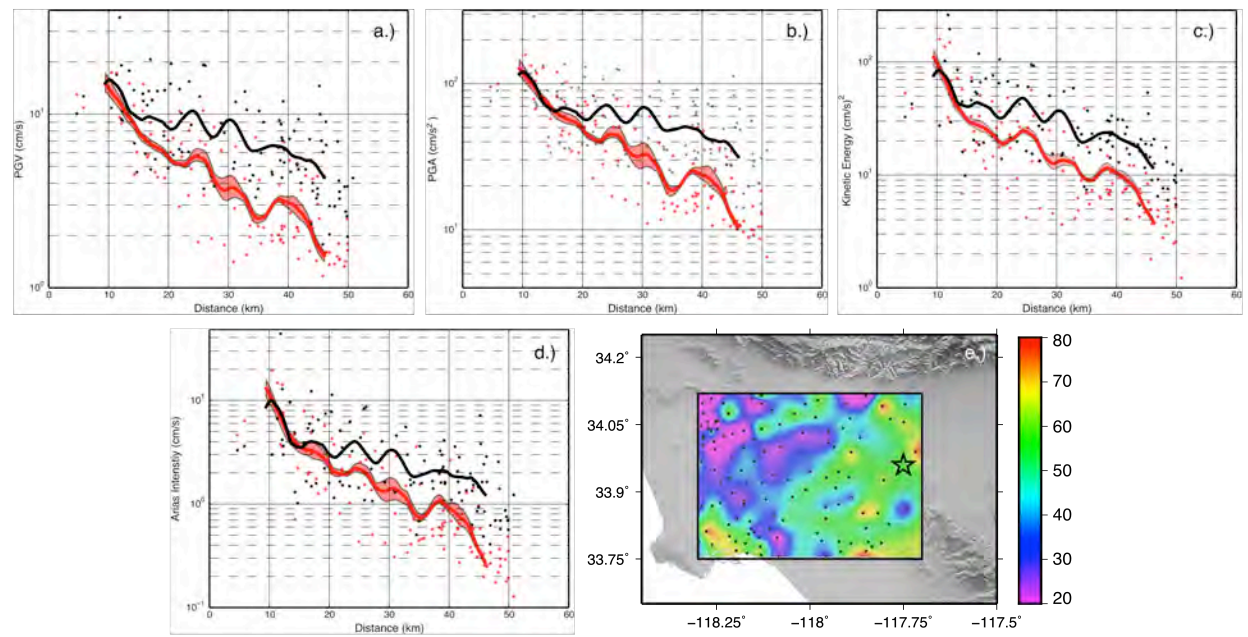
Figure 14.
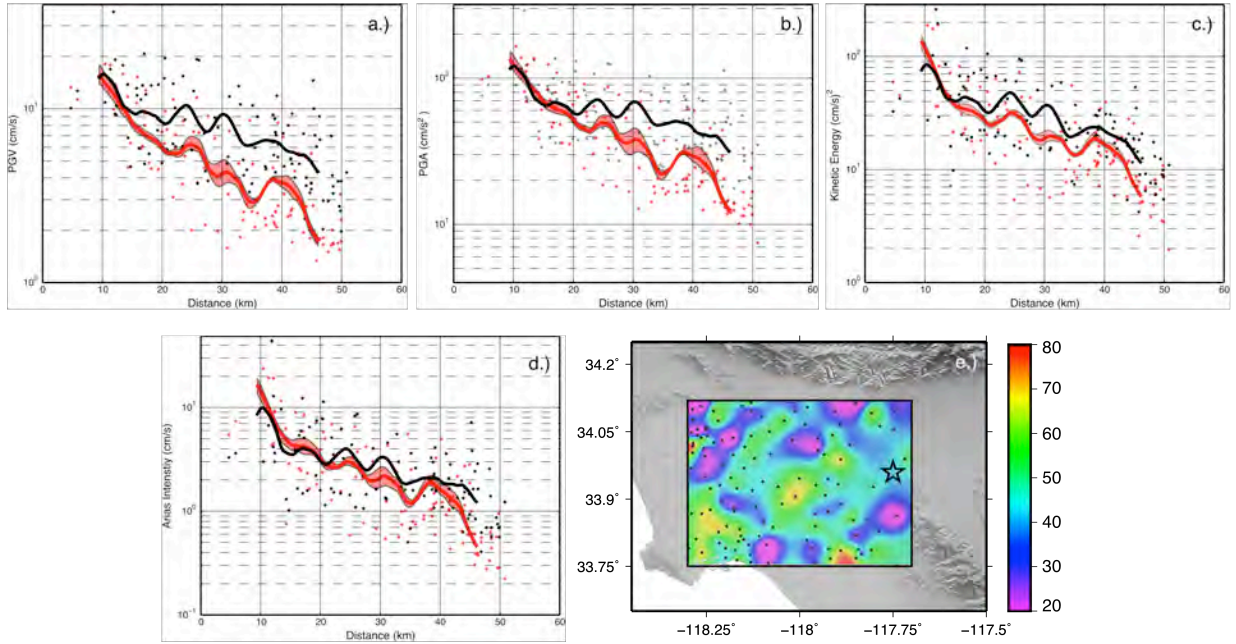
Figure 15.
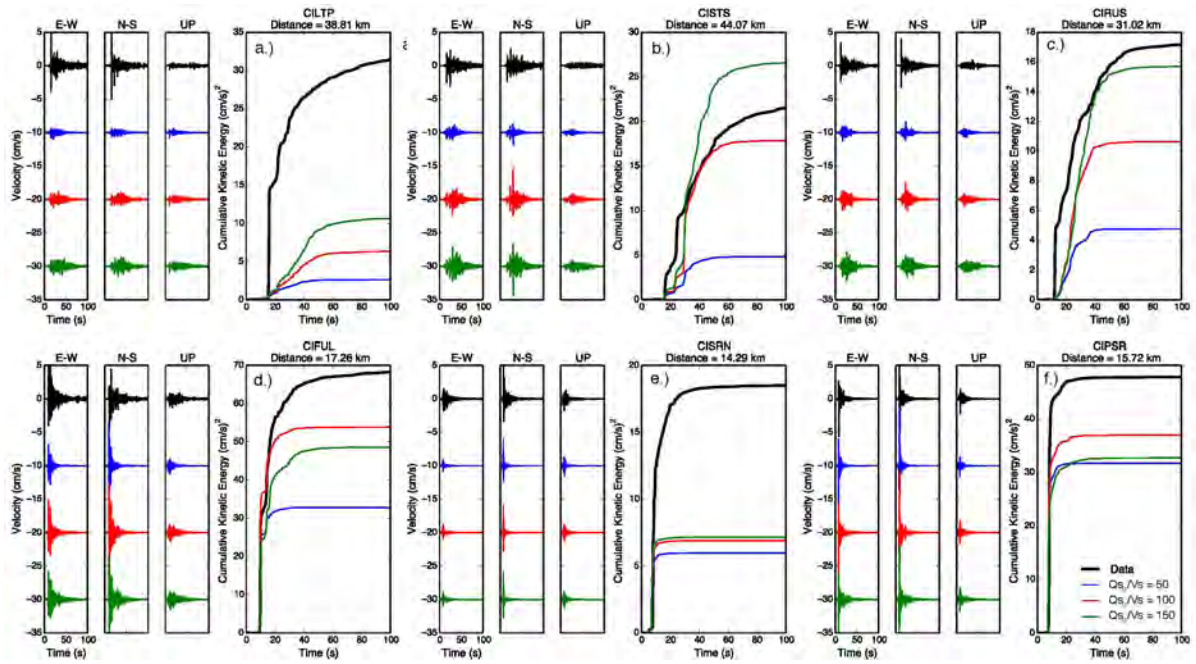
Figure 16.

Figure 17.

Figure 18.

Figure 19.

Figure 20.

Figure 21.